FIG. 1A.


```
        10         20         30         40         50
1234567890 1234567890 1234567890 1234567890 1234567890
TGGCCACTCC CTCTCTGCGC GCTCGCTCGC TCACTGAGGC CGGGCGACCA      50
AAGGTCGCCC GACGCCCGGG CTTTGCCCGG GCGGCCTCAG TGAGCGAGCG     100
AGCGCGCAGA GAGGGAGTGG CCAACTCCAT CACTAGGGGT TCCTCAGATC     150
TCTTTCTAAG TAAACAGTAC ATGAACCTTT ACCCCGTTGC TCGGCAACGG     200
CCTGGTCTGT GCCAAGTGTT TGCTGACGCA ACCCCCACTG GCTGGGGCTT     250
GGCCATAGGC CATCAGCGCA TGCGGATCTC AGTGTGGTTT TGCAAGAGGA     300
AGCAAAAAGC CTCTCCACCC AGGCCTGGAA TGTTTCCACC CAATGTCGAG     350
CAGTGTGGTT TTGCAAGAGG AAGCAAAAAG CCTCTCCACC CAGGCCTGGA     400
CTCGAGAGCT TCGACCACCA TGCAAATAGA GCTCTCCACC TGCTTCTTTC     450
                    M etGlnIleGl uLeuSerThr CysPhePheL
TGTGCCTTTT GCGATTCTGC TTTAGTGCCA CCAGAAGATA CTACCTGGGT     500
euCysLeuLe uArgPheCys PheSerAlaT hrArgArgTy rTyrLeuGly
GCAGTGGAAC TGTCATGGGA GTATATGCAA AGTGATCTCG GTGAGCTGCC     550
AlaValGluL euSerTrpAs pTyrMetGln SerAspLeuG lyGluLeuPr
TGTGGACGCA AGATTTCCTC CTAGAGTGCC AAAATCTTTT CCATTCAACA     600
oValAspAla ArgPheProP roArgValPr oLysSerPhe ProPheAsnT
CCTCAGTCGT GTACAAAAAG ACTCTGTTTG TAGAATTCAC GGTTCACCTT     650
hrSerValVa lTyrLysLys ThrLeuPheV alGluPheTh rValHisLeu
TTCAACATCG CTAAGCCAAG GCCACCCTGG ATGGGTCTGC TAGGTCCTAC     700
PheAsnIleA laLysProAr gProProTrp MetGlyLeuL euGlyProTh
CATCCAGGCT GAGGTTTATG ATACAGTGGT CATTACACTT AAGAACATGG     750
rIleGlnAla GluValTyrA spThrValVa lIleThrLeu LysAsnMetA
CTTCCCATCC TGTCAGTCTT CATGCTGTTG GTGTATCCTA CTGGAAAGCT     800
laSerHisPr oValSerLeu HisAlaValG lyValSerTy rTrpLysAla
TCTGAGGGAG CTGAATATGA TGATCAGACC AGTCAAAGGG AGAAAGAAGA     850
SerGluGlyA laGluTyrAs pAspGlnThr SerGlnArgG luLysGluAs
TGATAAAGTC TTCCCTGGTG CAAGCCATAC ATATGTCTGG CAGGTCCTGA     900
pAspLysVal PheProGlyG lySerHisTh rTyrValTrp GlnValLeuL
AAGAGAATGG TCCAATGGCC TCTGACCCAC TGTGCCTTAC CTACTCATAT     950
ysGluAsnGl yProMetAla SerAspProL euCysLeuTh rTyrSerTyr
```

FIG. 1B.


```
          10         20         30         40         50
1234567890 1234567890 1234567890 1234567890 1234567890
CTTTCTCATG TGGACCTGGT AAAAGACTTG AATTCAGGCC TCATTGGAGC  1000
LeuSerHisV alAspLeuVa lLysAspLeu AsnSerGlyL euIleGlyAl
CCTACTAGTA TGTAGAGAAG GGAGTCTGGC CAAGGAAAAG ACACAGACCT  1050
aleuLeuVal CysArgGluG lySerLeuAl aLysGluLys ThrGlnThrL
TGCACAAATT TATACTACTT TTTGCTGTAT TTGATGAAGG GAAAAGTTGG  1100
euHisLysPh eIleLeuLeu PheAlaValP heAspGluGl yLysSerTrp
CACTCAGAAA CAAAGAACTC CTTGATGCAG GATAGGGATG CTGCATCTGC  1150
HisSerGluT hrLysAsnSe rLeuMetGln AspArgAspA laAlaSerAl
TCGGGCCTGG CCTAAAATGC ACACAGTCAA TGGTTATGTA AACAGGTCTC  1200
aArgAlaTrp ProLysMetH isThrValAs nGlyTyrVal AsnArgSerL
TGCCAGGTCT GATTGGATGC CACAGGAAAT CAGTCTATTG GCATGTGATT  1250
euProGlyLe uIleGlyCys HisArgLysS erValTyrTr pHisValIle
GGAATGGGCA CCACTCCTGA AGTGCACTCA ATATTCCTCG AAGGTCACAC  1300
GlyMetGlyT hrThrProGl uValHisSer IlePheLeuG luGlyHisTh
ATTTCTTGTG AGGAACCATC GCCAGGCGTC CTTGGAAATC TCGCCAATAA  1350
rPheLeuVal ArgAsnHisA rgGlnAlaSe rLeuGluIle SerProIleT
CTTTCCTTAC TGCTCAAACA CTCTTGATGG ACCTTGGACA GTTTCTACTG  1400
hrPheLeuTh rAlaGlnThr LeuLeuMetA spLeuGlyGl nPheLeuLeu
TTTTGTCATA TCTCTTCCCA CCAACATGAT GGCATGGAAG CTTATGTCAA  1450
PheCysHisI leSerSerHi sGlnHisAsp GlyMetGluA laTyrValLy
AGTAGACAGC TGTCCAGAGG AACCCCAACT ACGAATGAAA AATAATGAAG  1500
sValAspSer CysProGluG luProGlnIle uArgMetLys AsnAsnGluG
AAGCGGAAGA CTATGATGAT GATCTTACTG ATTCTGAAAT GGATGTGGTC  1550
luAlaGluAs pTyrAspAsp AspLeuThrA spSerGluMe tAspValVal
AGGTTTGATG ATGACAACTC TCCTTCCTTT ATCCAAATTC GCTCAGTTGC  1600
ArgPheAspA spAspAsnSe rProSerPhe IleGlnIleA rgSerValAl
CAAGAAGCAT CCTAAAACTT GGGTACATTA CATTGCTGCT GAAGAGGAGG  1650
aLysLysHis ProLysThrT rpValHisTy rIleAlaAla GluGluGluA
ACTGGGACTA TGCTCCCTTA GTCCTGGCCC CCGATGACAG AAGTTATAAA  1700
spTrpAspTy rAlaProLeu ValLeuAlaP roAspAspAr gSerTyrLys
AGTCAATATT TGAACAATGG CCCTCAGCGG ATTGGTAGGA AGTACAAAAA  1750
SerGlnTyrL euAsnAsnGl yProGlnArg IleGlyArgL ysTyrLysLy
AGTCCGATTT ATGGCATACA CAGATGAAAC CTTTAAGACT CGTGAAGCTA  1800
sValArgPhe MetAlaTyrT hrAspGluTh rPheLysThr ArgGluAlaI
TTCAGCATGA ATCAGGAATC TTGGGACCTT TACTTTATGG GGAAGTTGGA  1850
leGlnHisGl uSerGlyIle LeuGlyProL euLeuTyrGl yGluValGly
GACACACTGT TGATTATATT TAAGAATCAA GCAAGCAGAC CATATAACAT  1900
AspThrLeuL euIleIlePh eLysAsnGln AlaSerArgP roTyrAsnIl
```

FIG. 1C.


```
        10         20         30         40         50
1234567890 1234567890 1234567890 1234567890 1234567890
CTACCCTCAC GGAATCACTG ATGTCCGTCC TTTGTATTCA AGGAGATTAC 1950
eTyrProHis GlyIleThrA spValArgPr oLeuTyrSer ArgArgLeuP
CAAAAGGTGT AAAACATTTG AAGGATTTTC CAATTCTGCC AGGAGAAATA 2000
roLysGlyVa lLysHisLeu LysAspPheP roIleLeuPr oGlyGluIle
TTCAAATATA AATGGACAGT GACTGTAGAA GATGGGCCAA CTAAATCAGA 2050
PheLysTyrL ysTrpThrVa lThrValGlu AspGlyProT hrLysSerAs
TCCTCGGTGC CTGACCCGCT ATTACTCTAG TTTCGTTAAT ATGGAGAGAG 2100
pProArgCys LeuThrArgT yrTyrSerSe rPheValAsn MetGluArgA
ATCTAGCTTC AGGACTCATT GGCCCTCTCC TCATCTGCTA CAAAGAATCT 2150
spLeuAlaSe rGlyLeuIle GlyProLeuL euIleCysTy rLysGluSer
GTAGATCAAA GAGGAAACCA GATAATGTCA GACAAGAGGA ATGTCATCCT 2200
ValAspGlnA rgGlyAsnGl nIleMetSer AspLysArgA snValIleLe
GTTTTCTGTA TTTGATGAGA ACCGAAGCTG GTACCTCACA GAGAATATAC 2250
uPheSerVal PheAspGluA snArgSerTr pTyrLeuThr GluAsnIleG
AACGCTTTCT CCCCAATCCA GCTGGAGTGC AGCTTGAGGA TCCAGAGTTC 2300
lnArgPheLe uProAsnPro AlaGlyValG lnLeuGluAs pProGluPhe
CAAGCCTCCA ACATCATGCA CAGCATCAAT GGCTATGTTT TTGATAGTTT 2350
GlnAlaSerA snIleMetHi sSerIleAsn GlyTyrValP heAspSerLe
GCAGTTGTCA GTTTGTTTGC ATGAGGTGGC ATACTGGTAC ATTCTAAGCA 2400
uGlnLeuSer ValCysLeuH isGluValAl aTyrTrpTyr IleLeuSerI
TTGGAGCACA GACTGACTTC CTTTCTGTCT TCTTCTCTGG ATATACCTTC 2450
leGlyAlaGl nThrAspPhe LeuSerValP hePheSerGl yTyrThrPhe
AAACACAAAA TGGTCTATGA AGACACACTC ACCCTATTCC CATTCTCAGG 2500
LysHisLysM etValTyrGl uAspThrLeu ThrLeuPheP roPheSerGl
AGAAACTGTC TTCATGTCGA TGGAAAACCC AGGTCTATGG ATTCTGGGGT 2550
yGluThrVal PheMetSerM etGluAsnPr oGlyLeuTrp IleLeuGlyC
GCCACAACTC AGACTTTCGG AACAGAGGCA TGACCGCCTT ACTGAAGGTT 2600
ysHisAsnSe rAspPheArg AsnArgGlyM etThrAlaLe uLeuLysVal
TCTAGTTGTG ACAAGAACAC TGGTGATTAT TACGAGGACA GTTATGAAGA 2650
SerSerCysA spLysAsnTh rGlyAspTyr TyrGluAspS erTyrGluAs
TATTTCAGCA TACTTGCTGA GTAAAAACAA TGCCATTGAA CCAAGAAGCT 2700
pIleSerAla TyrLeuLeuS erLysAsnAs nAlaIleGlu ProArgSerP
TCTCCCAGAA TTCAAGACAC CCTAGCACTA GGCAAAAGCA ATTTAATGCC 2750
heSerGlnAs nSerArgHis ProSerThrA rgGlnLysGl nPheAsnAla
ACCCCACCAG TCTTGAAACG CCATCAACGG GAAATAACTC GTACTACTCT 2800
ThrProProV alLeuLysAr gHisGlnArg GluIleThrA rgThrThrLe
TCAGTCAGAT CAAGAGGAAA TTGACTATGA TGATACCATA TCAGTTGAAA 2850
uGlnSerAsp GlnGluGluI leAspTyrAs pAspThrIle SerValGluM
```

FIG. 1D.


```
        10         20         30         40         50
1234567890 1234567890 1234567890 1234567890 1234567890
TGAAGAAGGA AGATTTTGAC ATTTATGATG AGGATGAAAA TCAGAGCCCC 2900
etLysLysGl uAspPheAsp IleTyrAspG luAspGluAs nGlnSerPro
CGCAGCTTTC AAAAGAAAAC ACGACACTAT TTTATTGCTG CAGTGGAGAG 2950
ArgSerPheG lnLysLysTh rArgHisTyr PheIleAlaA laValGluAr
GCTCTGGGAT TATGGGATGA GTAGCTCCCC ACATGTTCTA AGAAACAGGG 3000
gLeuTrpAsp TyrGlyMetS erSerSerPr oHisValLeu ArgAsnArgA
CTCAGAGTGG CAGTGTCCCT CAGTTCAAGA AAGTTGTTTT CCAGGAATTT 3050
laGlnSerGl ySerValPro GlnPheLysl ysValValPh eGlnGluPhe
ACTGATGGCT CCTTTACTCA GCCCTTATAC CGTGGAGAAC TAAATGAACA 3100
ThrAspGlyS erPheThrGl nProLeuTyr ArgGlyGluL euAsnGluHi
TTTGGGACTC CTGGGGCCAT ATATAAGAGC AGAAGTTGAA GATAATATCA 3150
sLeuGlyLeu LeuGlyProT yrIleArgAl aGluValGlu AspAsnIleM
TGGTAACTTT CAGAAATCAG GCCTCTGGTC CCTATTCCTT CTATTCTAGC 3200
etValThrPh eArgAsnGln AlaSerArgP roTyrSerPh eTyrSerSer
CTTATTTCTT ATGAGGAAGA TCAGAGGCAA GGAGCAGAAC CTAGAAAAAA 3250
LeuIleSerT yrGluGluAs pGlnArgGln GlyAlaGluP roArgLysAs
CTTTGTCAAG CCTAATGAAA CCAAAACTTA CTTTTGGAAA GTGCAACATC 3300
nPheValLys ProAsnGluT hrLysThrTy rPheTrpLys ValGlnHisH
ATATGGCACC CACTAAAGAT GAGTTTGACT GCAAAGCCTG GGCTTATTTC 3350
isMetAlaPr oThrLysAsp GluPheAspC ysLysAlaTr pAlaTyrPhe
TCTGATGTTG ACCTGGAAAA AGATGTGCAC TCAGGCCTGA TTGGACCCCT 3400
SerAspValA spLeuGluLy sAspValHis SerGlyLeuI leGlyProLe
TCTGGTCTGC CACACTAACA CACTGAACCC TGCTCATGGG AGACAAGTGA 3450
uLeuValCys HisThrAsnT hrLeuAsnPr oAlaHisGly ArgGlnValT
CAGTACAGGA ATTTGCTCTG TTTTTCACCA TCTTTGATGA GACCAAAAGC 3500
hrValGlnGl uPheAlaLeu PhePheThrI lePheAspGl uThrLysSer
TGGTACTTCA CTGAAAATAT GGAAAGAAAC TGCAGGGCTC CCTGCAATAT 3550
TrpTyrPheT hrGluAsnMe tGluArgAsn CysArgAlaP roCysAsnIl
CCAGATGGAA GATCCCACTT TTAAAGAGAA TTATGGCTTC CATGCAATCA 3600
eGlnMetGlu AspProThrP heLysGluAs nTyrArgPhe HisAlaIleA
ATGGCTACAT AATGGATACA CTACCTGGCT TAGTAATGGC TCAGGATCAA 3650
snGlyTyrIl eMetAspThr LeuProGlyL euValMetAl aGlnAspGln
AGGATTCGAT GGTATCTGCT CAGCATGGGC AGCAATGAAA ACATCCATTC 3700
ArgIleArgT rpTyrLeuLe uSerMetGly SerAsnGluA snIleHisSe
TATTCATTTC AGTGGACATG TGTTCACTGT ACGAAAAAAA GAGGAGTATA 3750
rIleHisPhe SerGlyHisV alPheThrVa lArgLysLys GluGluTyrL
AAATGGCACT GTACAATCTC TATCCAGGTG TTTTTGAGAC AGTGGAAATG 3800
ysMetAlaLe uTyrAsnLeu TyrProGlyV alPheGluTh rValGluMet
```

FIG. 1E.

10	20	30	40	50	
1234567890	1234567890	1234567890	1234567890	1234567890	
TTACCATCCA	AAGCTGGAAT	TTGGCGGGTG	GAATGCCTTA	TTGGCGAGCA	3850
LeuProSerL	ysAlaGlyIl	eTrpArgVal	GluCysLeuI	leGlyGluHi	
TCTACATGCT	GGGATGAGCA	CACTTTTTCT	GGTGTACAGC	AATAAGTGTC	3900
sLeuHisAla	GlyMetSerT	hrLeuPheLe	uValTyrSer	AsnLysCysG	
AGACTCCCCT	GGGAATGGCT	TCTGGACACA	TTAGAGATTT	TCAGATTACA	3950
lnThrProLe	uGlyMetAla	SerGlyHisI	leArgAspPh	eGlnIleThr	
GCTTCAGGAC	AATATGGACA	GTGGGCCCCA	AAGCTGGCCA	GACTTCATTA	4000
AlaSerGlyG	lnTyrGlyGl	nTrpAlaPro	LysLeuAlaA	rgLeuHisTy	
TTCCGGATCA	ATCAATGCCT	GGAGCACCAA	GGAGCCCTTT	TCTTGGATCA	4050
rSerGlySer	IleAsnAlaT	rpSerThrLy	sGluProPhe	SerTrpIleL	
AGGTGGATCT	GTTGGCACCA	ATGATTATTC	ACGGCATCAA	GACCCAGGGT	4100
ysValAspLe	uLeuAlaPro	MetIleIleH	isGlyIleLy	sThrGlnGly	
GCCCGTCAGA	AGTTCTCCAG	CCTCTACATC	TCTCAGTTTA	TCATCATGTA	4150
AlaArgGlnL	ysPheSerSe	rLeuTyrIle	SerGlnPheI	leIleMetTy	
TAGTCTTGAT	GGGAAGAAGT	GGCAGACTTA	TCGAGGAAAT	TCCACTGGAA	4200
rSerLeuAsp	GlyLysLysT	rpGlnThrTy	rArgGlyAsn	SerThrGlyT	
CCTTAATGGT	CTTCTTTGGC	AATGTGGATT	CATCTGGGAT	AAAACACAAT	4250
hrLeuMetVa	lPhePheGly	AsnValAspS	erSerGlyIl	eLysHisAsn	
ATTTTTAACC	CTCCAATTAT	TGCTCGATAC	ATCCGTTTGC	ACCCAACTCA	4300
IlePheAsnP	roProIleIl	eAlaArgTyr	IleArgLeuH	isProThrHi	
TTATAGCATT	CGCAGCACTC	TTCGCATGGA	GTTGATGGGC	TGTGATTTAA	4350
sTyrSerIle	ArgSerThrL	euArgMetGl	uLeuMetGly	CysAspLeuA	
ATAGTTGCAG	CATGCCATTG	GGAATGGAGA	GTAAAGCAAT	ATCAGATGCA	4400
snSerCysSe	rMetProLeu	GlyMetGluS	erLysAlaIl	eSerAspAla	
CAGATTACTG	CTTCATCCTA	CTTTACCAAT	ATGTTTGCCA	CCTGGTCTCC	4450
GlnIleThrA	laSerSerTy	rPheThrAsn	MetPheAlaT	hrTrpSerPr	
TTCAAAAGCT	CGACTTCACC	TCCAAGGGAG	GAGTAATGCC	TGGAGACCTC	4500
oSerLysAla	ArgLeuHisL	euGlnGlyAr	gSerAsnAla	TrpArgProG	
AGGTGAATAA	TCCAAAAGAG	TGGCTGCAAG	TGGACTTCCA	GAAGACAATG	4550
lnValAsnAs	nProLysGlu	TrpLeuGlnV	alAspPheGl	nLysThrMet	
AAAGTCACAG	GAGTAACTAC	TCAGGGAGTA	AAATCTCTGC	TTACCAGCAT	4600
LysValThrG	lyValThrTh	rGlnGlyVal	LysSerLeuL	euThrSerMe	
GTATGTGAAG	GAGTTCCTCA	TCTCCAGCAG	TCAAGATGGC	CATCAGTGGA	4650
tTyrValLys	GluPheLeuI	leSerSerSe	rGlnAspGly	HisGlnTrpT	
CTCTCTTTTT	TCAGAATGGC	AAAGTAAAGG	TTTTTCAGGG	AAATCAAGAC	4700
hrLeuPhePh	eGlnAsnGly	LysValLysV	alPheGlnGl	yAsnGlnAsp	
TCCTTCACAC	CTGTGGTGAA	CTCTCTAGAC	CCACCGTTAC	TGACTCGCTA	4750
SerPheThrP	roValValAs	nSerLeuAsp	ProProLeuL	euThrArgTy	

FIG. 1F.


```
        10        20        30        40        50
1234567890 1234567890 1234567890 1234567890 1234567890
CCTTCGAATT CACCCCCAGA GTTGGGTGCA CCAGATTGCC CTGAGGATGG  4800
rLeuArgIle HisProGlnS erTrpValHi sGlnIleAla LeuArgMetG
AGGTTCTGGG CTGCGAGGCA CAGGACCTCT ACTGACTCGA GCGAGTTCTT  4850
luValLeuGl yCysGluAla GlnAspLeuT yr...
CTGAGGGGAT CGGCAATAAA AAGACAGAAT AAAACGCACG GGTGTTGGGT  4900

CGTTTGTTCG GATCCAGATC TAGGAACCCC TAGTGATGGA GTTGGCCACT  4950

CCCTCTCTGC GCGCTCGCTC GCTCACTGAG GCCGCCCGGG CAAAGCCCGG  5000

GCGTCGGGCG ACCTTTGGTC GCCCGGCCTC AGTGAGCGAG CGAGCGCGCA  5050

GAGAGGGAGT GGCCAACCCC CCCCCCCCCC CCCCTGCAGC CCAGCTGCAT  5100

TAATGAATCG GCCAACGCGC GGGGAGAGGC GGTTTGCGTA TTGGGCGCTC  5150

TTCCGCTTCC TCGCTCACTG ACTCGCTGCG CTCGGTCGTT CGGCTGCGGC  5200

GAGCGGTATC AGCTCACTCA AAGGCGGTAA TACGGTTATC CACAGAATCA  5250

GGGGATAACG CAGGAAAGAA CATGTGAGCA AAAGGCCAGC AAAAGGCCAG  5300

GAACCGTAAA AAGGCCGCGT TGCTGGCGTT TTTCCATAGG CTCCGCCCCC  5350

CTGACGAGCA TCACAAAAAT CGACGCTCAA GTCAGAGGTG GCGAAACCCG  5400

ACAGGACTAT AAAGATACCA GGCGTTTCCC CCTGGAAGCT CCCTCGTGCG  5450

CTCTCCTGTT CCGACCCTGC CGCTTACCGG ATACCTGTCC GCCTTTCTCC  5500

CTTCGGGAAG CGTGGCGCTT TCTCAATGCT CACGCTGTAG GTATCTCAGT  5550

TCGGTGTAGG TCGTTCGCTC CAAGCTGGGC TGTGTGCACG AACCCCCCGT  5600

TCAGCCCGAC CGCTGCGCCT TATCCGGTAA CTATCGTCTT GAGTCCAACC  5650

CGGTAAGACA CGACTTATCG CCACTGGCAG CAGCCACTGG TAACAGGATT  5700
```

FIG. 1G.


```
        10         20         30         40         50
1234567890 1234567890 1234567890 1234567890 1234567890
AGCAGAGCGA GGTATGTAGG CGGTGCTACA GAGTTCTTGA AGTGGTGGCC  5750

TAACTACGGC TACACTAGAA GGACAGTATT TGGTATCTGC GCTCTGCTGA  5800

AGCCAGTTAC CTTCGGAAAA AGAGTTGGTA GCTCTTGATC CGGCAAACAA  5850

ACCACCGCTG GTAGCGGTGG TTTTTTTGTT TGCAAGCAGC AGATTACGCG  5900

CAGAAAAAAA GGATCTCAAG AAGATCCTTT GATCTTTTCT ACGGGGTCTG  5950

ACGCTCAGTG GAACGAAAAC TCACGTTAAG GGATTTTGGT CATGAGATTA  6000

TCAAAAAGGA TCTTCACCTA GATCCTTTTA AATTAAAAAT GAAGTTTTAA  6050

ATCAATCTAA AGTATATATG AGTAAACTTG GTCTGACAGT TACCAATGCT  6100
                                                  ylGelIreS
TAATCAGTGA GGCACCTATC TCAGCGATCT GTCTATTTCG TTCATCCATA  6150
ueL...siHo rPlaV...gr AueLreSgrA psAelIulGn sAteMprTue
GTTGCCTGAC TCCCCGTCGT GTAGATAACT ACGATACGGG AGGGCTTACC  6200
LnlGgrAlaV ylGgrAgrAr hTreSueL.. .reSlaVorP orPreSlaVt
ATCTGGCCCC AGTGCTGCAA TGATACCGCG AGACCCACGC TCACCGGCTC  6250
eMnlGylGpr TsiHnlGueL reSlaValAu eLylGlaVre SlaVorPulG
CAGATTTATC AGCAATAAAC CAGCCAGCCG GAAGGGCCGA GCGCAGAAGT  6300
ueLnsAelIu eLueLueLyl GalAueLgrA ehPorPgrAa lAsyCehPsi
GGTCCTGCAA CTTTATCCGC CTCCATCCAG TCTATTAATT GTTGCCGGGA  6350
HpsAnlGueL syLelIgrAg rAprTylGrh T......nsA nsAylGorPu
AGCTAGAGTA AGTAGTTCGC CAGTTAATAG TTTGCGCAAC GTTGTTGCCA  6400
eL...ueLue LryTnsAalA ueL...ryTn sAalAsyCgr AnlGnlGprT
TTGCTACAGG CATCGTGGTG TCACGCTCGT CGTTTGGTAT GGCTTCATTC  6450
nlG...ueLs yCgrAorPrh TlaVreSrhT rhTnlGryTo rPsyLteM..
AGCTCCGGTT CCCAACGATC AAGGCGAGTT ACATGATCCC CCATGTTGTG  6500
.reSgrAnsA ylGlaVelIu eLalAueL.. .teMelIylG prTrhTrhTs
CAAAAAAGCG GTTAGCTCCT TCGGTCCTCC GATCGTTGTC AGAAGTAAGT  6550
yCehPueLor P...reSgrA grApsAulGr eSgrAnlG.. .ehPryTrhT
TGGCCGCAGT GTTATCACTC ATGGTTATGG CAGCACTGCA TAATTCTCTT  6600
orPgrAueLr hTelIlaV.. .orP...orP ueLlaValAr yTnsAulG..
ACTGTCATGC CATCCGTAAG ATGCTTTTCT GTGACTGGTG AGTACTCAAC  6650
.nlG...alA teMgrAueLe lIreSsyLnl GreSnlGsiH rhTreSueLp
```

FIG. 1H.

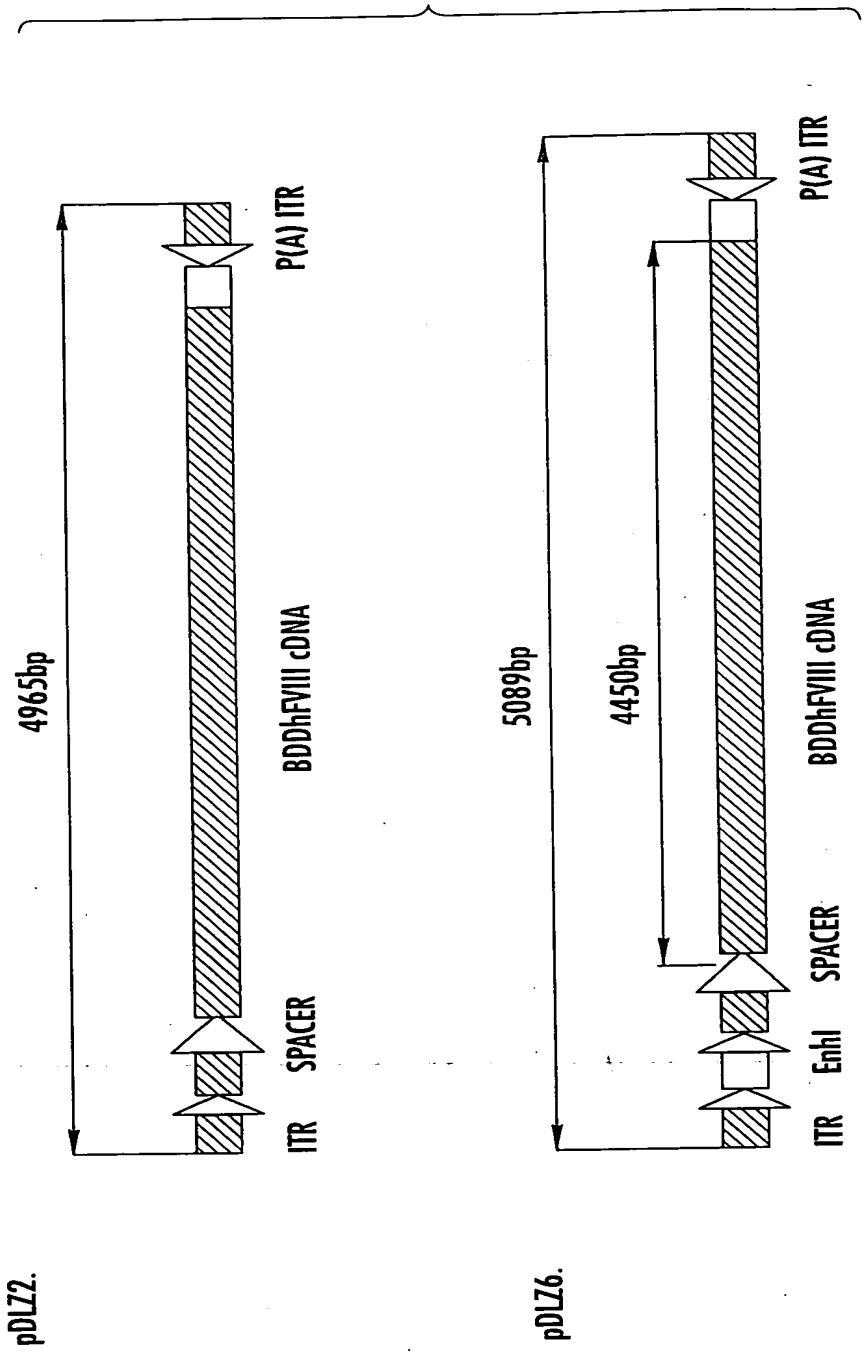

```
          10         20         30         40         50
1234567890 1234567890 1234567890 1234567890 1234567890
CAAGTCATTC TGAGAATAGT GTATGCGGCG ACCGAGTTGC TCTTGCCCGG  6700
rTrhTteMgr AueLelIrhT ryTalAalAl aVreSnsAre SsyLylGorP
CGTCAATACG GGATAATACC GCGCCACATA GCAGAACTTT AAAAGTGCTC  6750
rhTueLlaVo rPryTryTgr AalAlaVryT syCehPsyLu eLueLalA..
ATCATTGGAA AACGTTCTTC GGGGCGAAAA CTCTCAAGGA TCTTACCGCT  6800
....nlGehP laVnsAsyLo rPalAehPla VgrAueLreS grAlaValAr
GTTGAGATCC AGTTCGATGT AACCCACTCG TGCACCCAAC TGATCTTCAG  6850
hTreSelIpr TnsAreSrhT laVprTulGs iHlaVprTre SelIsyLueL
CATCTTTTAC TTTCACCAGC GTTTCTGGGT GAGCAAAAAC AGGAAGGCAA  6900
teMsyL...s yL...prTgr AsyLnlGrhT ueLueLehPu eLehPalAeh
AATGCCGCAA AAAAGGGAAT AAGGGCGACA CGGAAATGTT GAATACTCAT  6950
PsiHgrAueL ehPorPehPu eLorPreSla VreSelInsA ehPlaV...l
ACTCTTCCTT TTTCAATATT ATTGAAGCAT TTATCAGGGT TATTGTCTCA  7000

TGAGCGGATA CATATTTGAA TGTATTTAGA AAAATAAACA AATAGGGGTT  7050

CCGCGCACAT TTCCCCGAAA AGTGCCACCT GACGTCTAAG AAACCATTAT  7100

TATCATGACA TTAACCTATA AAAATAGGCG TATCACGAGG CCCTTTCGTC  7150

TCGCGCGTTT CGGTGATGAC GGTGAAAACC TCTGACACAT GCAGCTCCCG  7200

GAGACGGTCA CAGCTTGTCT GTAAGCGGAT GCCGGGAGCA GACAAGCCCG  7250

TCAGGGCGCG TCAGCGGGTG TTGGCGGGTG TCGGGGCTGG CTTAACTATG  7300

CGGCATCAGA GCAGATTGTA CTGAGAGTGC ACCATATGCG GTGTGAAATA  7350

CCGCACAGAT GCGTAAGGAG AAAATACCGC ATCAGGAAAT TGTAAACGTT  7400

AATATTTTGT TAAAATTCGC GTTAAATTTT TGTTAAATCA GCTCATTTTT  7450

TAACCAATAG GCCGAAATCG GCAAAATCCC TTATAAATCA AAAGAATAGA  7500

CCGAGATAGG GTTGAGTGTT GTTCCAGTTT GGAACAAGAG TCCACTATTA  7550

AAGAACGTGG ACTCCAACGT CAAAGGGCGA AAAACCGTCT ATCAGGGCGA  7600
```

FIG. 1I.


```
          10         20         30         40         50
1234567890 1234567890 1234567890 1234567890 1234567890
TGGCCCACTA CGTGAACCAT CACCCTAATC AAGTTTTTTG GGGTCGAGGT  7650

GCCGTAAAGC ACTAAATCGG AACCCTAAAG GGAGCCCCCG ATTTAGAGCT  7700

TGACGGGGAA AGCCGGCGAA CGTGGCGAGA AAGGAAGGGA AGAAAGCGAA  7750

AGGAGCGGGC GCTAGGGCGC TGGCAAGTGT AGCGGTCACG CTGCGCGTAA  7800

CCACCACACC CGCCGCGCTT AATGCGCCGC TACAGGGCGC GTCGCGCCAT  7850

TCGCCATTCA GGCTACGCAA CTGTTGGGAA GGGCGATCGG TGCGGGCCTC  7900

TTCGCTATTA CGCCAGCTGG CTGCAGGGGG GGGGGGGGGG GGGT        7944
```

Title: *Adeno-Associated Virus Vectors Encoding Factor VIII and Methods of Using the Same*
Inventor(s): Walsh et al.
Application No: Not Assigned
Atty Dkt No: 035052/270239

FIG. 2.



Title: *Adeno-Associated Virus Vectors Encoding Factor VIII and Methods of Using the Same*
Inventor(s): Walsh et al.
Application No: Not Assigned
Atty Dkt No: 035052/270239

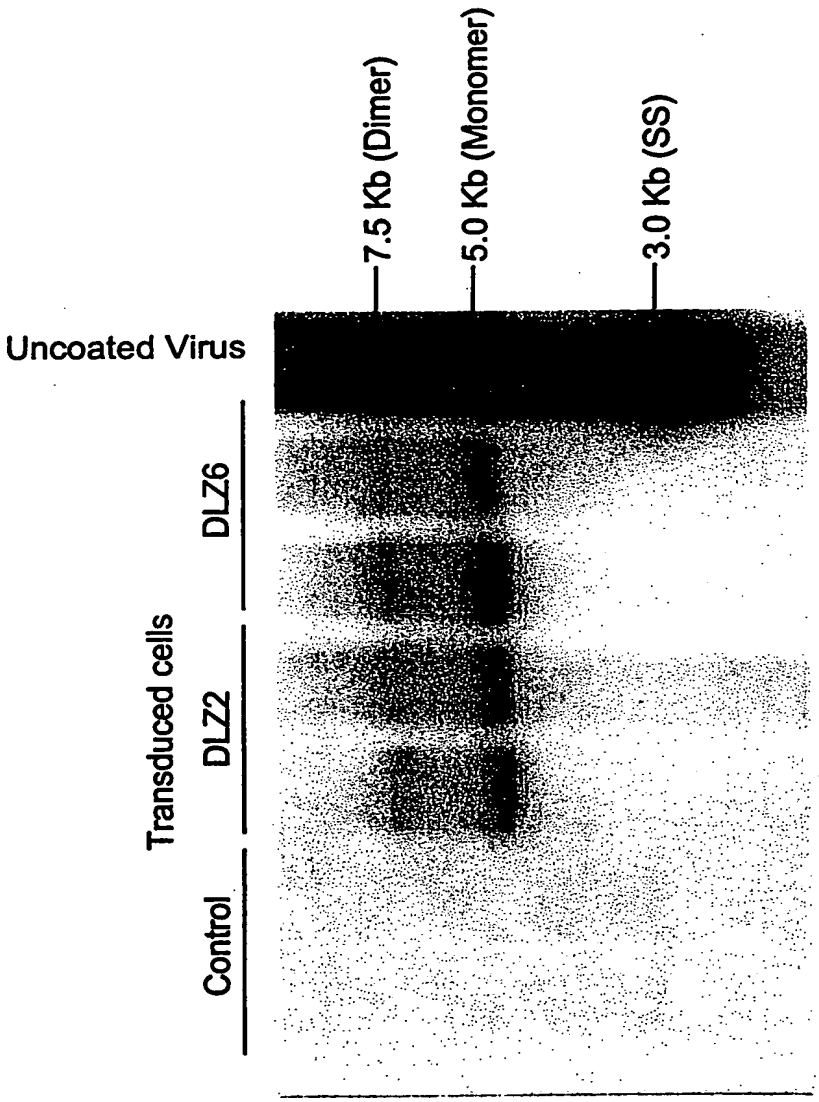
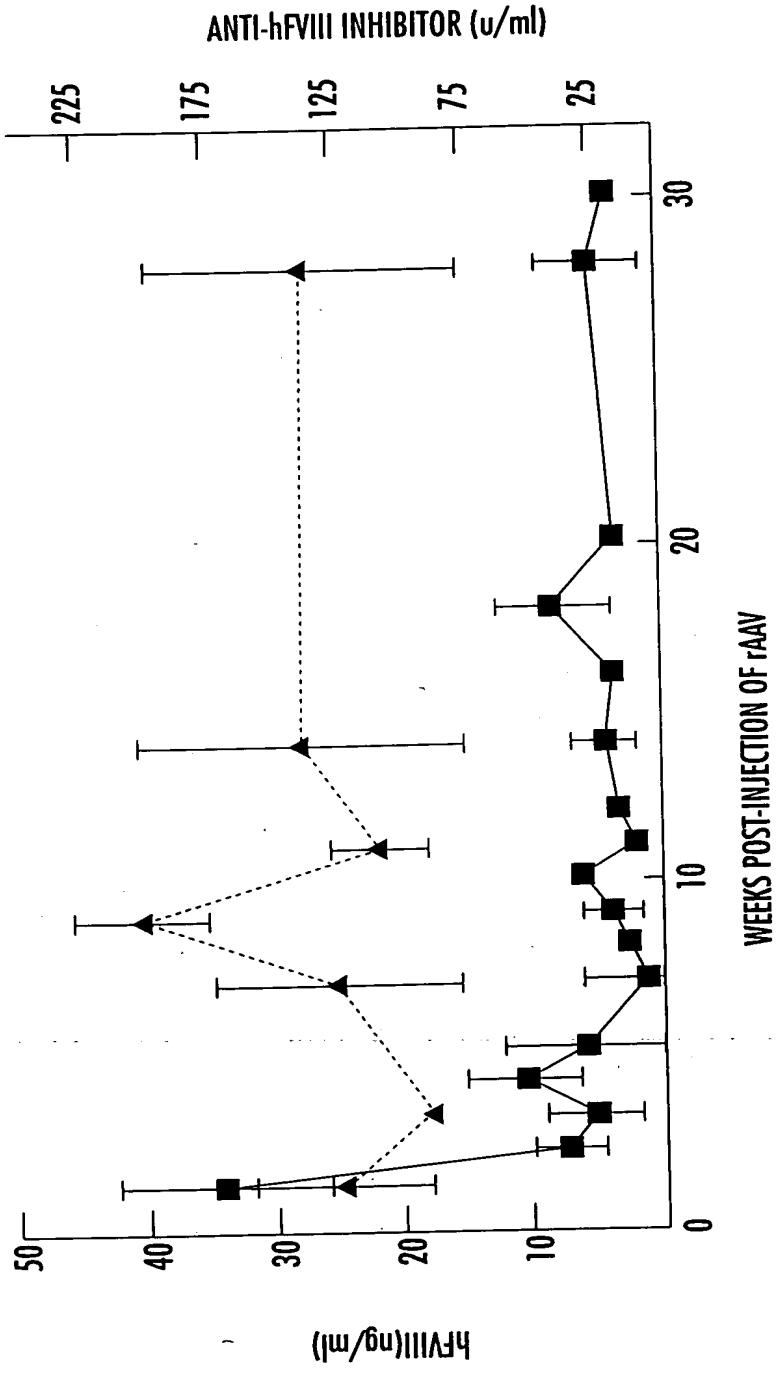


FIG. 3.

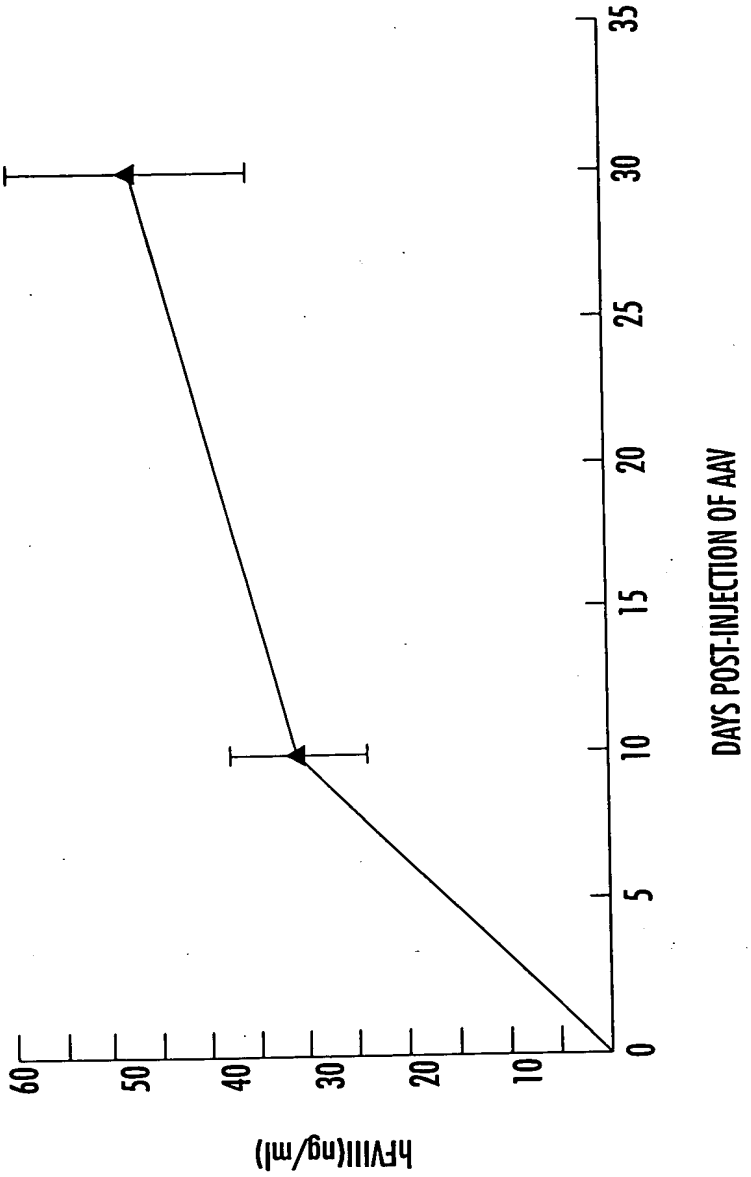
Title: *Adeno-Associated Virus Vectors Encoding Factor VIII and Methods of Using the Same*
Inventor(s): Walsh et al.
Application No: Not Assigned
Atty Dkt No: 035052/270239

FIG. 4.A.



Title: *Adeno-Associated Virus Vectors Encoding Factor VIII and Methods of Using the Same*
Inventor(s): Walsh et al.
Application No: Not Assigned
Atty Dkt No: 035052/270239

FIG. 4.B.



Title: *Adeno-Associated Virus Vectors Encoding Factor VIII and Methods of Using the Same*
Inventor(s): Walsh et al.
Application No: Not Assigned
Atty Dkt No: 035052/270239

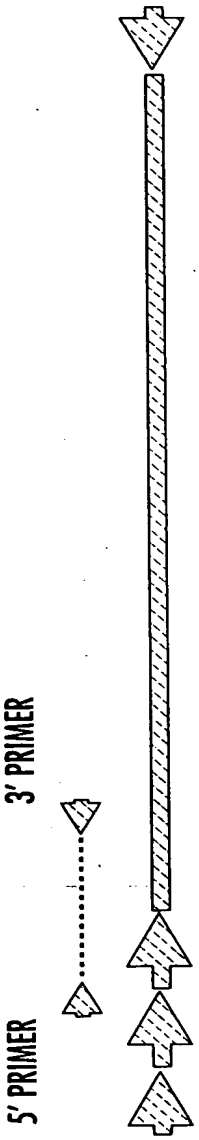


FIG. 5.A.

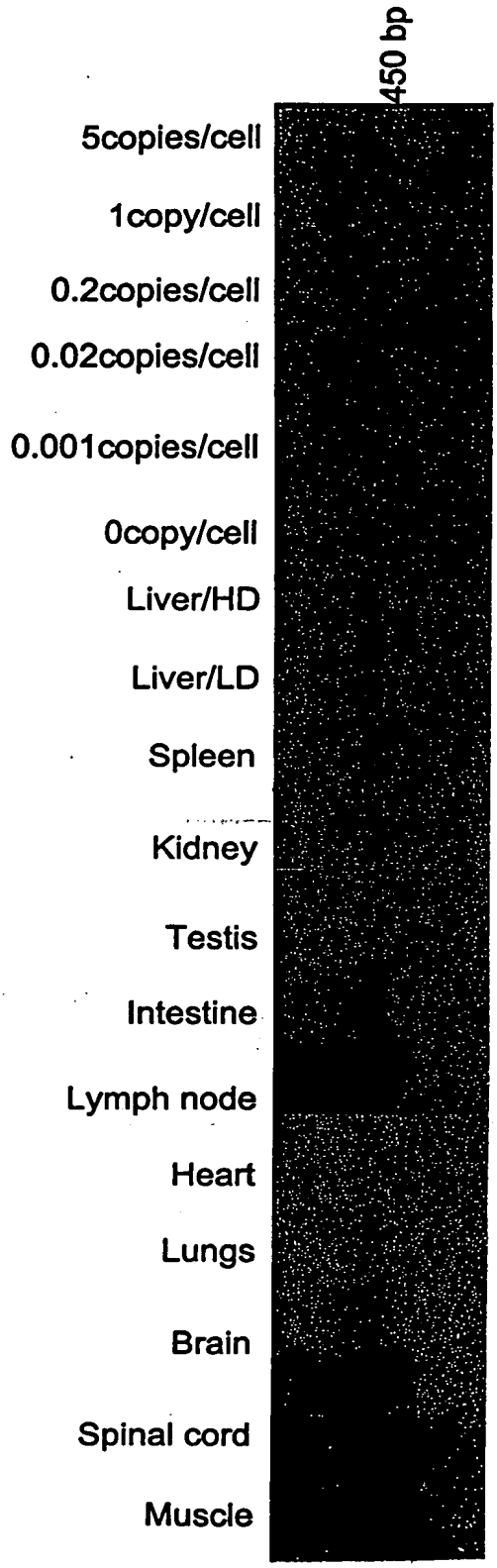


FIG. 5.B.

Title: *Adeno-Associated Virus Vectors Encoding Factor VIII and Methods of Using the Same*
Inventor(s): Walsh et al.
Application No: Not Assigned
Atty Dkt No: 035052/270239

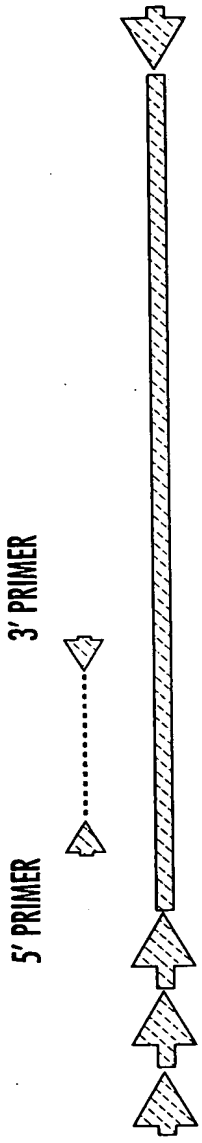


FIG. 5.C.

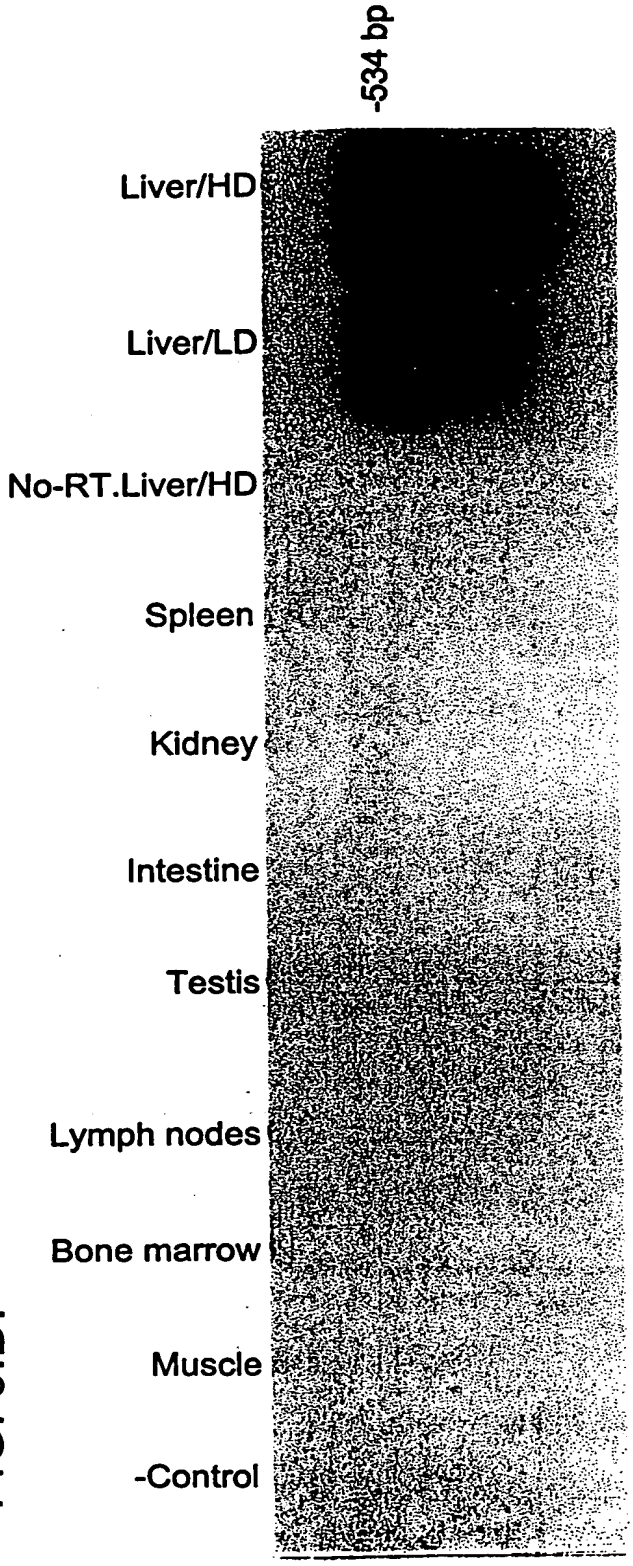


FIG. 5.D.

FIG. 5.E.

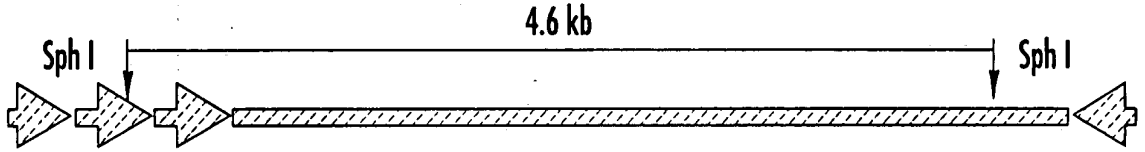

4.6 kb
Sph I
Sph I

FIG. 5.F.

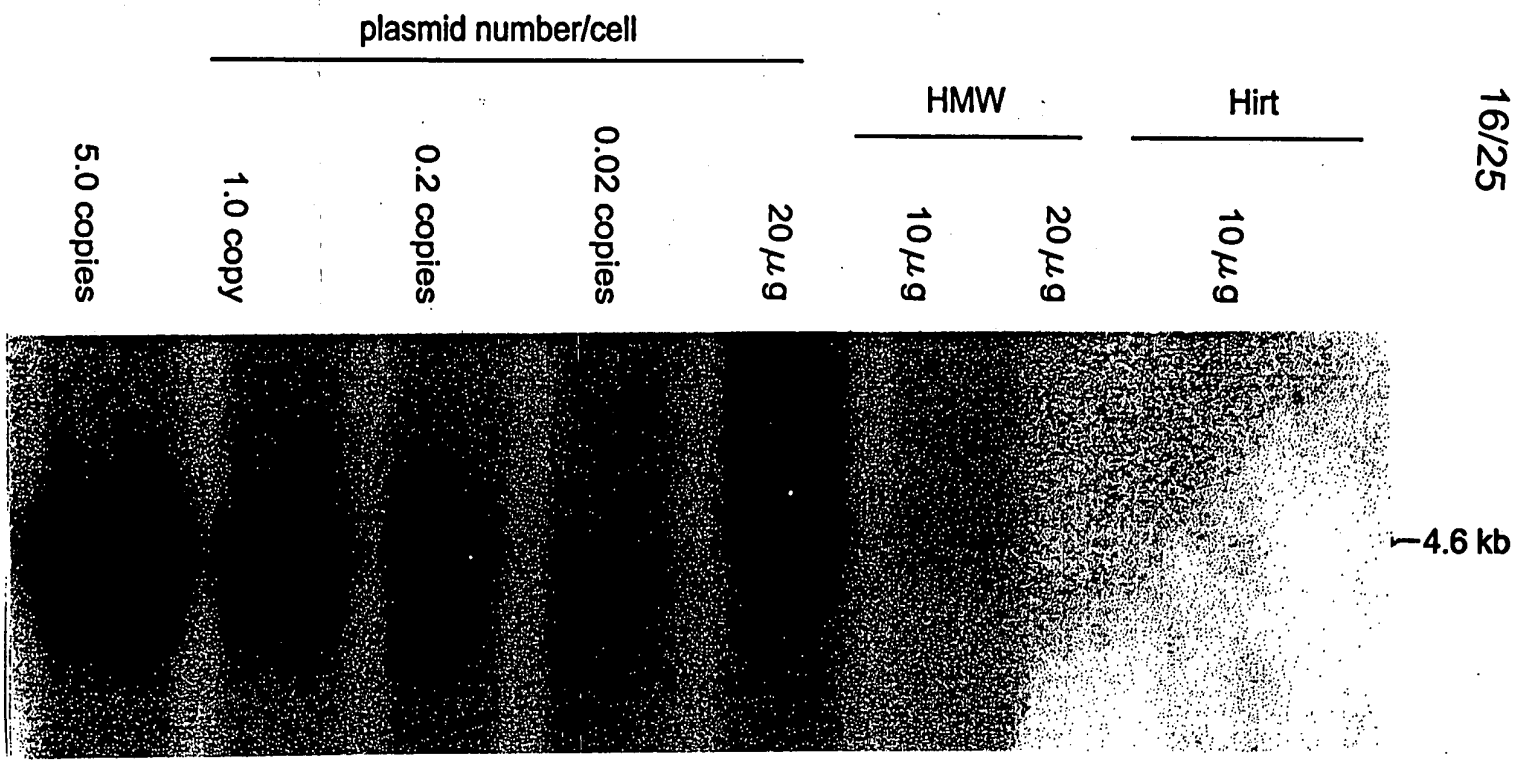

plasmid number/cell
5.0 copies
1.0 copy
0.2 copies
0.02 copies
20 μg
HMW
10 μg
20 μg
Hirt
10 μg
4.6 kb

FIG. 6.A.


```
          10         20         30         40         50
1234567890 1234567890 1234567890 1234567890 1234567890
TGGCCACTCC CTCTCTGCGC GCTCGCTCGC TCACTGAGGC CGGGCGACCA   50

AAGGTCGCCC GACGCCCGGG CTTTGCCCGG GCGGCCTCAG TGAGCGAGCG  100

AGCGCGCAGA GAGGGAGTGG CCAACTCCAT CACTAGGGGT TCCTCAGATC  150

TCTTTCTAAG TAAACAGTAC ATGAACCTTT ACCCCGTTGC TCGGCAACGG  200

CCTGGTCTGT GCCAAGTGTT TGCTGACGCA ACCCCCACTG GCTGGGGCTT  250

GGCCATAGGC CATCAGCGCA TGCGGATCTC AGTGTGGTTT TGCAAGAGGA  300

AGCAAAAAGC CTCTCCACCC AGGCCTGGAA TGTTTCCACC CAATGTCGAG  350

CAGTGTGGTT TTGCAAGAGG AAGCAAAAAG CCTCTCCACC CAGGCCTGGA  400

CTCGACCTCG AGAGTACTTC TAGAAATACG AGCCATGCAA GTAGAGCTCT  450
                                    MetGln ValGluLeuT
ACACCTGCTG CTTTCTGTGC CTTTTGCCCT TCAGCCTTAG TGCCACCAGA  500
yrThrCysCy sPheLeuCys LeuLeuProP heSerLeuSe rAlaThrArg
AAATACTACC TCGGTGCAGT GGAACTGTCC TGGGACTATA TGCAAAGTGA  550
LysTyrTyrL euGlyAlaVa lGluLeuSer TrpAspTyrM etGlnSerAs
CCTGCTCAGT GCGCTGCACG CGGATACAAG CTTTTCTTCC AGGGTGCCAG  600
pLeuLeuSer AlaLeuHisA laAspThrSe rPheSerSer ArgValProG
GATCTTTGCC ACTCACCACG TCAGTCACGT ACAGAAAGAC TGTGTTTGTA  650
lySerLeuPr oLeuThrThr SerValThrT yrArgLysTh rValPheVal
GAGTTTACAG ATGACCTTTT CAACATTGCC AAGCCCAGGC CACCGTGGAT  700
GluPheThrA spAspLeuPh eAsnIleAla LysProArgP roProTrpMe
GGGCCTGCTG GGTCCTACCA TCCAGGCTGA GGTTTATGAC ACAGTGGTCA  750
tGlyLeuLeu GlyProThrI leGlnAlaGl uValTyrAsp ThrValValI
TTGTCCTTAA GAACATGGCT TCTCATCCTG TCAGCCTTCA CGCTGTTGGT  800
leValLeuLy sAsnMetAla SerHisProV alSerLeuHi sAlaValGly
GTATCCTATT GGAAAGCTTC TGAAGGTGCT GAGTATGAGG ATCAGACCAG  850
ValSerTyrT rpLysAlaSe rGluGlyAla GluTyrGluA spGlnThrSe
CCAAAAGGAG AAGGAAGATG ATAATGTCAT TCCTGGTGAA AGCCATACCT  900
rGlnLysGlu LysGluAspA spAsnValIl eProGlyGlu SerHisThrT
ATGTCTGGCA GGTCCTGAAA GAGAATGGCC CAATGGCCTC TGATCCACCA  950
yrValTrpGl nValLeuLys GluAsnGlyP roMetAlaSe rAspProPro
```

FIG. 6.B.


```
        10         20         30         40         50
1234567890 1234567890 1234567890 1234567890 1234567890
TGTCTCACCT ACTCATATTT TTCACACGTG GACCTGGTGA AAGACCTGAA 1000
CysLeuThrT yrSerTyrPh eSerHisVal AspLeuValL ysAspLeuAs
TTCAGGCCTC ATTGGAGCCC TGCTGGTTTG CAAAGAAGGG AGTCTGGCCA 1050
nSerGlyLeu IleGlyAlaL euLeuValCy sLysGluGly SerLeuAlaL
AAGAAAGGAC ACAGACCTTG CAGGAATTTG TCCTACTTTT TGCTGTATTT 1100
ysGluArgTh rGlnThrLeu GlnGluPheV alLeuLeuPh eAlaValPhe
GATGAAGGGA AAAGTTGGCA CTCAGAAACA AATGCGTCTT TGACACAGGC 1150
AspGluGlyL ysSerTrpHi sSerGluThr AsnAlaSerL euThrGlnAl
TGAGGCCCAG CATGAGCTGC ACACCATCAA TGGCTATGTA AACAGGTCTC 1200
aGluAlaGln HisGluLeuH isThrIleAs nGlyTyrVal AsnArgSerL
TGCCAGGTCT TACTGTGTGT CACAAGAGAT CAGTCTATTG GCATGTGATT 1250
euProGlyLe uThrValCys HisLysArgS erValTyrTr pHisValIle
GGAATGGGCA CCACCCCCGA AGTGCACTCA ATTTTTCTCG AAGGTCACAC 1300
GlyMetGlyT hrThrProGl uValHisSer IlePheLeuG luGlyHisTh
ATTTCTTGTG AGGAACCACC GCCAGGCCTC CTTGGAGATC TCACCAATTA 1350
rPheLeuVal ArgAsnHisA rgGlnAlaSe rLeuGluIle SerProIleT
CTTTCCTTAC TGCTCAGACA TTCCTGATGG ACCTTGGCCA GTTTCTACTG 1400
hrPheLeuTh rAlaGlnThr PheLeuMetA spLeuGlyGl nPheLeuLeu
TTTTGTCATA TCCCTTCCCA TCAACATGAT GGTATGGAAG CTTATGTCAA 1450
PheCysHisI leProSerHi sGlnHisAsp GlyMetGluA laTyrValLy
AGTAGATAGC TGCCCAGAGG AACCCCAGCT GCGCATGAAA AATAATGAAG 1500
sValAspSer CysProGluG luProGlnLe uArgMetLys AsnAsnGluA
ATAAAGATTA TGATGATGGT CTTTATGATT CTGACATGGA CGTAGTTAGC 1550
spLysAspTy rAspAspGly LeuTyrAspS erAspMetAs pValValSer
TTTGATGACG ACAGCTCTTC TCCCTTTATC CAAATCCGCT CAGTTGCCAA 1600
PheAspAspA spSerSerSe rProPheIle GlnIleArgS erValAlaLy
GAAGCATCCT AAAACTTGGG TCCACTATAT TGCTGCTGAG GAGGAGGACT 1650
sLysHisPro LysThrTrpV alHisTyrIl eAlaAlaGlu GluGluAspT
GGGACTATGC TCCCTCAGGC CCCACCCCCA ATGATAGAAG TCATAAAAAT 1700
rpAspTyrAl aProSerGly ProThrProA snAspArgSe rHisLysAsn
CTGTATTTGA ACAATGGTCC TCAGCGGATT GGTAAGAAGT ACAAAAAAGT 1750
LeuTyrLeuA snAsnGlyPr oGlnArgIle GlyLysLysT yrLysLysVa
CCGATTTGTG GCATACACAG ATGAGACATT TAAGACTCGT GAAGCTATTC 1800
lArgPheVal AlaTyrThrA spGluThrPh eLysThrArg GluAlaIleG
AGTATGAATC AGGAATCCTG GGACCTTTAC TTTATGGAGA AGTTGGAGAC 1850
lnTyrGluSe rGlyIleLeu GlyProLeuL euTyrGlyGl uValGlyAsp
ACACTGCTGA TTATATTTAA GAATCAAGCC AGCCGGCCAT ATAACATCTA 1900
ThrLeuLeuI leIlePheLy sAsnGlnAla SerArgProT yrAsnIleTy
```

FIG. 6.C.


```
        10         20         30         40         50
1234567890 1234567890 1234567890 1234567890 1234567890
CCCTCATGGG ATCAATTATG TCACTCCTCT GCACACAGGG AGATTGCCAA 1950
rProHisGly IleAsnTyrV alThrProLe uHisThrGly ArgLeuProL
AAGGTGTGAA ACATTTGAAA GATATGCCAA TTCTGCCGGG AGAGATATTC 2000
ysGlyValLy sHisLeuLys AspMetProI leLeuProGl yGluIlePhe
AAGTATAAAT GGACAGTGAC CGTAGAAGAT GGACCAACTA AATCAGATCC 2050
LysTyrLysT rpThrValTh rValGluAsp GlyProThrL ysSerAspPr
TCGGTGCCTG ACCCGATATT ACTCAAGCTT CATTAATCTG GAGAGAGATC 2100
oArgCysLeu ThrArgTyrT yrSerSerPh eIleAsnLeu GluArgAspL
TAGCTTCAGG ACTCATTGGC CCTCTTCTCA TCTGCTACAA AGAATCTGTA 2150
euAlaSerGl yLeuIleGly ProLeuLeuI leCysTyrLy sGluSerVal
GATCAAAGAG GAAACCAGAT GATGTCAGAC AAGAGAAATG TCATCCTGTT 2200
AspGlnArgG lyAsnGlnMe tMetSerAsp LysArgAsnV alIleLeuPh
TTCTGTATTT GATGAGAATC GAAGCTGGTA CCTCACAGAG AATATGCAGC 2250
eSerValPhe AspGluAsnA rgSerTrpTy rLeuThrGlu AsnMetGlnA
GCTTCCTCCC CAATGCAGAT GTAGTGCAGC CCCATGACCC AGAGTTCCAA 2300
rgPheLeuPr oAsnAlaAsp ValValGlnP roHisAspPr oGluPheGln
CTCTCTAACA TCATGCACAG CATCAATGGC TATGTTTTTG ACAACTTGCA 2350
LeuSerAsnI leMetHisSe rIleAsnGly TyrValPheA spAsnLeuGl
GCTGTCAGTT TGTTTGCATG AGGTGGCGTA CTGGTACATT CTAAGTGTTG 2400
nLeuSerVal CysLeuHisG luValAlaTy rTrpTyrIle LeuSerValG
GAGCACAAAC TGACTTCCTG TCTGTCTTCT TCTCTGGATA TACCTTCAAA 2450
lyAlaGlnTh rAspPheLeu SerValPheP heSerGlyTy rThrPheLys
CACAAAATGG TCTATGAAGA CACACTTACC CTCTTCCCAT TCTCAGGAGA 2500
HisLysMetV alTyrGluAs pThrLeuThr LeuPheProP heSerGlyGl
AACTGTCTTC ATGTCAATGG AAAACCCAGG TCTGTGGGTT CTGGGGTGCC 2550
uThrValPhe MetSerMetG luAsnProGl yLeuTrpVal LeuGlyCysH
ACAACTCAGA CTTTCGGAAC AGAGGCATGA CAGCCTTACT GAAGGTTTCT 2600
isAsnSerAs pPheArgAsn ArgGlyMetT hrAlaLeuLe uLysValSer
AGTTGTAACA GGAACATTGA TGATTATTAT GAGGACACAT ACGAAGATAT 2650
SerCysAsnA rgAsnIleAs pAspTyrTyr GluAspThrT yrGluAspIl
TCCAACTCCC CTGCTAAATG AAAACAATGT AATTAAACCT AGAAGCTTCT 2700
eProThrPro LeuLeuAsnG luAsnAsnVa lIleLysPro ArgSerPheS
CCCAGAATTC AAGGCACCCT AGCACTAAGG AAAAGCAATT GAAAATGAAG 2750
erGlnAsnSe rArgHisPro SerThrLysG luLysGlnLe uLysMetLys
AGAGAAGATT TTGACATCTA CGGCGACTAT GAAAATCAGG GCCTCCGCAG 2800
ArgGluAspP heAspIleTy rGlyAspTyr GluAsnGlnG lyLeuArgSe
CTTTCAAAAG AAAACACGAC ACTATTTCAT TGCTGCAGTG GAGCGTCTCT 2850
rPheGlnLys LysThrArgH isTyrPheIl eAlaAlaVal GluArgLeuT
```

FIG. 6.D.


```
          10         20         30         40         50
1234567890 1234567890 1234567890 1234567890 1234567890
GGGATTATGG GATGAGTAGA TCTCCCCATA TACTAAGAAA CAGGGCTCAA 2900
rpAspTyrGl yMetSerArg SerProHisI leLeuArgAs nArgAlaGln
AGTGGGGATG TCCAGCAGTT CAAGAAGGTG GTTTTCCAGG AATTTACTGA 2950
SerGlyAspV alGlnGlnPh eLysLysVal ValPheGlnG luPheThrAs
TGGATCCTTT ACTCAGCCCT TATACCGTGG AGAACTGAAT GAACACTTGG 3000
pGlySerPhe ThrGlnProL euTyrArgGl yGluLeuAsn GluHisLeuG
GACTCTTGGG GCCATATATA AGAGCAGAAG TTGAAGACAA TATCGTGGTA 3050
lyLeuLeuGl yProTyrIle ArgAlaGluV alGluAspAs nIleValVal
ACTTTCAAAA ACCAGGCCTC TCGTCCCTAC TCCTTCTATT CTAGTCTTAT 3100
ThrPheLysA snGlnAlaSe rArgProTyr SerPheTyrS erSerLeuIl
TTCTTATGAC GAAGATGAGG GACAAGGAGC AGAACCTAGA AGAAAGTTTG 3150
eSerTyrAsp GluAspGluG lyGlnGlyAl aGluProArg ArgLysPheV
TCAACCCTAA TGAAACCAAA ATTTACTTTT GGAAAGTGCA GCATCATATG 3200
alAsnProAs nGluThrLys IleTyrPheT rpLysValGl nHisHisMet
GCACCCACTA AAGATGAGTT TGACTGCAAA GCCTGGGCTT ATTTTTCTGA 3250
AlaProThrL ysAspGluPh eAspCysLys AlaTrpAlaT yrPheSerAs
TGTTGATTTG GAGAAAGATG TGCACTCAGG CTTGATTGGA CCCCTTCTGA 3300
pValAspLeu GluLysAspV alHisSerGl yLeuIleGly ProLeuLeuI
TCTGCCGCAG TAACACACTG AACCCTGCTC ATGGGAGACA AGTGACAGTG 3350
leCysArgSe rAsnThrLeu AsnProAlaH isGlyArgGl nValThrVal
CAGGAGTTTG CCCTGGTTTT CACTATATTC GATGAGACTA AGAGCTGGTA 3400
GlnGluPheA laLeuValPh eThrIlePhe AspGluThrL ysSerTrpTy
CTTCACTGAA AACCTGGAAA GGAACTGTAG AGCTCCCTGC AATGTCCAGA 3450
rPheThrGlu AsnLeuGluA rgAsnCysAr gAlaProCys AsnValGlnL
AGGAGGACCC TACTCTAAAA GAAAACTTCC GCTTCCATGC AATCAACGGC 3500
ysGluAspPr oThrLeuLys GluAsnPheA rgPheHisAl aIleAsnGly
TATGTGAAGG ATACACTCCC TGGCTTAGTA ATGGCTCAGG ATCAAAAGGT 3550
TyrValLysA spThrLeuPr oGlyLeuVal MetAlaGlnA spGlnLysVa
TCGATGGTAT CTGCTCAGCA TGGGCAGCAA CGAAAACATT CATTCCATTC 3600
lArgTrpTyr LeuLeuSerM etGlySerAs nGluAsnIle HisSerIleH
ACTTCAGTGG ACATGTGTTC ACTGTACGGA AAAAAGAGGA ATATAAAATG 3650
isPheSerGl yHisValPhe ThrValArgL ysLysGluGl uTyrLysMet
GCAGTCTACA ACCTCTATCC AGGTGTTTTT GAGACTGTGG AAATGCTACC 3700
AlaValTyrA snLeuTyrPr oGlyValPhe GluThrValG luMetLeuPr
ATCCCAAGTT GGAATCTGGC GGATAGAATG CCTTATCGGC GAGCACCTGC 3750
oSerGlnVal GlyIleTrpA rgIleGluCy sLeuIleGly GluHisLeuG
AAGCCGGGAT GAGCACTCTG TTTCTGGTGT ACAGCAAGAA GTGTCAGACT 3800
lnAlaGlyMe tSerThrLeu PheLeuValT yrSerLysLy sCysGlnThr
```

FIG. 6.E.


```
        10        20        30        40        50
1234567890 1234567890 1234567890 1234567890 1234567890
CCACTGGGGA TGGCTTCCGG ACACATTAGA GATTTTCAGA TTACAGCTTC  3850
ProLeuGlyM etAlaSerGl yHisIleArg AspPheGlnI leThrAlaSe
AGGACAATAT GGACAGTGGG CCCCAAAGCT GGCCAGACTT CATTATTCCG  3900
rGlyGlnTyr GlyGlnTrpA laProLysLe uAlaArgLeu HisTyrSerG
GATCAATCAA TGCCTGGAGC ACCAAGGATC CCTTTTCCTG GATCAAGGTG  3950
lySerIleAs nAlaTrpSer ThrLysAspP roPheSerTr pIleLysVal
GATCTCTTGG CACCGATGAT TATTCACGGC ATCATGACCC AGGGGGCCCG  4000
AspLeuLeuA laProMetIl eIleHisGly IleMetThrG lnGlyAlaAr
CCAGAAGTTC TCCAGCCTCT ACGTGTCTCA GTTTATCATC ATGTACAGTC  4050
gGlnLysPhe SerSerLeuT yrValSerGl nPheIleIle MetTyrSerL
TGGATGGCAA CAAGTGGCAC AGTTACCGAG GGAATTCCAC GGGGACCTTA  4100
euAspGlyAs nLysTrpHis SerTyrArgG lyAsnSerTh rGlyThrLeu
ATGGTCTTCT TTGGCAACGT GGATTCATCT GGGATCAAAC ACAATATTTT  4150
MetValPheP heGlyAsnVa lAspSerSer GlyIleLysH isAsnIlePh
TAACCCTCCG ATTATTGCTC AGTACATCCG TTTGCACCCA ACCCATTACA  4200
eAsnProPro IleIleAlaG lnTyrIleAr gLeuHisPro ThrHisTyrS
GCATCCGCAG CACTCTTCGC ATGGAGCTCT TGGGCTGTGA CTTCAACAGT  4250
erIleArgSe rThrLeuArg MetGluLeuL euGlyCysAs pPheAsnSer
TGCAGCATGC CGCTGGGGAT GGAGAGTAAA GCAATATCAG ATGCTCAGAT  4300
CysSerMetP roLeuGlyMe tGluSerLys AlaIleSerA spAlaGlnIl
CACTGCCTCG TCCTACCTAA GCAGTATGCT TGCCACTTGG TCTCCTTCCC  4350
eThrAlaSer SerTyrLeuS erSerMetIe uAlaThrTrp SerProSerG
AAGCCCGGCT GCACCTGCAG GGCAGGACTA ATGCCTGGAG ACCTCAGGCA  4400
lnAlaArgLe uHisLeuGln GlyArgThrA snAlaTrpAr gProGlnAla
AATAACCCAA AAGAGTGGCT GCAAGTGGAC TTCCGGAAGA CCATGAAAGT  4450
AsnAsnProL ysGluTrpLe uGlnValAsp PheArgLysT hrMetLysVa
CACAGGAATA ACCACCCAGG GGGTGAAATC TCTCCTCATC AGCATGTATG  4500
lThrGlyIle ThrThrGlnG lyValLysSe rLeuLeuIle SerMetTyrV
TGAAGGAGTT CCTCATCTCC AGTAGTCAAG ATGGCCATAA CTGGACTCTG  4550
alLysGluPh eLeuIleSer SerSerGlnA spGlyHisAs nTrpThrLeu
TTTCTTCAGA ATGGCAAAGT CAAGGTCTTC CAGGGAAACC GGGACTCCTC  4600
PheLeuGlnA snGlyLysVa lLysValPhe GlnGlyAsnA rgAspSerSe
CACGCCTGTG CGGAACCGTC TCGAACCCCC GCTGGTGGCT CGCTACGTGC  4650
rThrProVal ArgAsnArgL euGluProPr oLeuValAla ArgTyrValA
GCCTGCACCC GCAGAGCTGG CGGCACCACA TCGCCCTGAG GCTGGAGGTC  4700
rgLeuHisPr oGlnSerTrp AlaHisHisI leAlaLeuAr gLeuGluVal
CTGGGCTGCG ACACCCAGCA GCCCGCCTGA CCCGCGCCTC TGCGGCCCTG  4750
LeuGlyCysA spThrGlnGl nProAla...
```

FIG. 6.F.

10 20 30 40 50
1234567890 1234567890 1234567890 1234567890 1234567890

TCTCCCCTGC CTCCCTGCCC TGTCCCCGCG GCTTCCCATC AAGCTTATCG 4800

ATACCGTCGA GCGAGTTCTT CTGAGGGGAT CGGCAATAAA AAGACAGAAT 4850

AAAACGCACG GGTGTTGGGT CGTTTGTTCG GATCCAGATC TAGGAACCCC 4900

TAGTGATGGA GTTGGCCACT CCCTCTCTGC GCGCTCGCTC GCTCACTGAG 4950

GCCGCCCGGG CAAAGCCCGG GCGTCGGGCG ACCTTTGGTC GCCCGGCCTC 5000

AGTGAGCGAG CGAGCGCGCA GAGAGGGAGT GGCCAACCCC CCCCCCCCCC 5050

CCCCTGCAGC CCAGCTGCAT TAATGAATCG GCCAACGCGC GGGGAGAGGC 5100

GGTTTGCGTA TTGGGCGCTC TTCCGCTTCC TCGCTCACTG ACTCGCTGCG 5150

CTCGGTCGTT CGGCTGCGGC GAGCGGTATC AGCTCACTCA AAGGCGGTAA 5200

TACGGTTATC CACAGAATCA GGGGATAACG CAGGAAAGAA CATGTGAGCA 5250

AAAGGCCAGC AAAAGGCCAG GAACCGTAAA AAGGCCGCGT TGCTGGCGTT 5300

TTTCCATAGG CTCCGCCCCC CTGACGAGCA TCACAAAAAT CGACGCTCAA 5350

GTCAGAGGTG GCGAAACCCG ACAGGACTAT AAAGATACCA GGCGTTTCCC 5400

CCTGGAAGCT CCCTCGTGCG CTCTCCTGTT CCGACCCTGC CGCTTACCGG 5450

ATACCTGTCC GCCTTTCTCC CTTCGGGAAG CGTGGCGCTT TCTCAATGCT 5500

CACGCTGTAG GTATCTCAGT TCGGTGTAGG TCGTTCGCTC CAAGCTGGGC 5550

TGTGTGCACG AACCCCCCGT TCAGCCCGAC CGCTGCGCCT TATCCGGTAA 5600

CTATCGTCTT GAGTCCAACC CGGTAAGACA CGACTTATCG CCACTGGCAG 5650

CAGCCACTGG TAACAGGATT AGCAGAGCGA GGTATGTAGG CGGTGCTACA 5700

FIG. 6.G.


```
          10         20         30         40         50
1234567890 1234567890 1234567890 1234567890 1234567890
GAGTTCTTGA AGTGGTGGCC TAACTACGGC TACACTAGAA GGACAGTATT  5750

TGGTATCTGC GCTCTGCTGA AGCCAGTTAC CTTCGGAAAA AGAGTTGGTA  5800

GCTCTTGATC CGGCAAACAA ACCACCGCTG GTAGCGGTGG TTTTTTTGTT  5850

TGCAAGCAGC AGATTACGCG CAGAAAAAAA GGATCTCAAG AAGATCCTTT  5900

GATCTTTTCT ACGGGGTCTG ACGCTCAGTG GAACGAAAAC TCACGTTAAG  5950

GGATTTTGGT CATGAGATTA TCAAAAAGGA TCTTCACCTA GATCCTTTTA  6000

AATTAAAAAT GAAGTTTTAA ATCAATCTAA AGTATATATG AGTAAACTTG  6050

GTCTGACAGT TACCAATGCT TAATCAGTGA GGCACCTATC TCAGCGATCT  6100
          ylGelIreS ueL...siHo rPlaV...gr AueLreSgrA
GTCTATTTCG TTCATCCATA GTTGCCTGAC TCCCCGTCGT GTAGATAACT  6150
psAelIulGn sAteMprTue LnlGgrAlaV ylGgrAgrAr hTreSueL..
ACGATACGGG AGGGCTTACC ATCTGGCCCC AGTGCTGCAA TGATACCGCG  6200
.reSlaVorP orPreSlaVt eMnlGylGpr TsiHnlGueL reSlaValAu
AGACCCACGC TCACCGGCTC CAGATTTATC AGCAATAAAC CAGCCAGCCG  6250
eLylGlaVre SlaVorPulG ueLnsAelIu eLueLueLyl GalAueLgrA
GAAGGGCCGA GCGCAGAAGT GGTCCTGCAA CTTTATCCGC CTCCATCCAG  6300
ehPorPgrAa lAsyCehPsi HpsAnlGueL syLelIgrAg rAprTylGrh
TCTATTAATT GTTGCCGGGA AGCTAGAGTA AGTAGTTCGC CAGTTAATAG  6350
T......nsA nsAylGorPu eL...ueLue LryTnsAalA ueL...ryTn
TTTGCGCAAC GTTGTTGCCA TTGCTACAGG CATCGTGGTG TCACGCTCGT  6400
sAalAsyCgr AnlGnlGprT nlG...ueLs yCgrAorPrh TlaVreSrhT
CGTTTGGTAT GGCTTCATTC AGCTCCGGTT CCCAACGATC AAGGCGAGTT  6450
rhTnlGryTo rPsyLteM.. .reSgrAnsA ylGlaVelIu eLalAueL..
ACATGATCCC CCATGTTGTG CAAAAAAGCG GTTAGCTCCT TCGGTCCTCC  6500
.teMelIylG prTrhTrhTs yCehPueLor P...reSgrA grApsAulGr
GATCGTTGTC AGAAGTAAGT TGGCCGCAGT GTTATCACTC ATGGTTATGG  6550
eSgrAnlG.. .ehPryTrhT orPgrAueLr hTelIlaV.. .orP...orP
CAGCACTGCA TAATTCTCTT ACTGTCATGC CATCCGTAAG ATGCTTTTCT  6600
ueLlaValAr yTnsAulG.. .nlG...alA teMgrAueLe lIreSsyLnl
GTGACTGGTG AGTACTCAAC CAAGTCATTC TGAGAATAGT GTATGCGGCG  6650
GreSnlGsiH rhTreSueLp rTrhTteMgr AueLelIrhT ryTalAalAl
```

FIG. 6.H.


```
          10         20         30         40         50
1234567890 1234567890 1234567890 1234567890 1234567890
ACCGAGTTGC TCTTGCCCGG CGTCAATACG GGATAATACC GCGCCACATA 6700
aVreSnsAre SsyLylGorP rhTueLlaVo rPryTryTgr AalAlaVryT
GCAGAACTTT AAAAGTGCTC ATCATTGGAA AACGTTCTTC GGGGCGAAAA 6750
syCehPsyLu eLueLalA.. ....nlGehP laVnsAsyLo rPalAehPla
CTCTCAAGGA TCTTACCGCT GTTGAGATCC AGTTCGATGT AACCCACTCG 6800
VgrAueLreS grAlaValAr hTreSelIpr TnsAreSrhT laVprTulGs
TGCACCCAAC TGATCTTCAG CATCTTTTAC TTTCACCAGC GTTTCTGGGT 6850
iHlaVprTre SelIsyLueL teMsyL...s yL...prTgr AsyLnlGrhT
GAGCAAAAAC AGGAAGGCAA AATGCCGCAA AAAAGGGAAT AAGGGCGACA 6900
ueLueLehPu eLehPalAeh PsiHgrAueL ehPorPehPu eLorPreSla
CGGAAATGTT GAATACTCAT ACTCTTCCTT TTTCAATATT ATTGAAGCAT 6950
VreSelInsA ehPlaV...l
TTATCAGGGT TATTGTCTCA TGAGCGGATA CATATTTGAA TGTATTTAGA 7000

AAAATAAACA AATAGGGGTT CCGCGCACAT TTCCCCGAAA AGTGCCACCT 7050

GACGTCTAAG AAACCATTAT TATCATGACA TTAACCTATA AAAATAGGCG 7100

TATCACGAGG CCCTTTCGTC TCGCGCGTTT CGGTGATGAC GGTGAAAACC 7150

TCTGACACAT GCAGCTCCCG GAGACGGTCA CAGCTTGTCT GTAAGCGGAT 7200

GCCGGGAGCA GACAAGCCCG TCAGGGCGCG TCAGCGGGTG TTGGCGGGTG 7250

TCGGGGCTGG CTTAACTATG CGGCATCAGA GCAGATTGTA CTGAGAGTGC 7300

ACCATATGCG GTGTGAAATA CCGCACAGAT GCGTAAGGAG AAAATACCGC 7350

ATCAGGAAAT TGTAAACGTT AATATTTTGT TAAAATTCGC GTTAAATTTT 7400

TGTTAAATCA GCTCATTTTT TAACCAATAG GCCGAAATCG GCAAAATCCC 7450

TTATAAATCA AAAGAATAGA CCGAGATAGG GTTGAGTGTT GTTCCAGTTT 7500

GGAACAAGAG TCCACTATTA AAGAACGTGG ACTCCAACGT CAAAGGGCGA 7550

AAAACCGTCT ATCAGGGCGA TGGCCCACTA CGTGAACCAT CACCCTAATC 7600
```

FIG. 6.I.

10 20 30 40 50
1234567890 1234567890 1234567890 1234567890 1234567890

AAGTTTTTTG GGGTCGAGGT GCCGTAAAGC ACTAAATCGG AACCCTAAAG 7650

GGAGCCCCCG ATTTAGAGCT TGACGGGGAA AGCCGGCGAA CGTGGCGAGA 7700

AAGGAAGGGA AGAAAGCGAA AGGAGCGGGC GCTAGGGCGC TGGCAAGTGT 7750

AGCGGTCACG CTGCGCGTAA CCACCACACC CGCCGCGCTT AATGCGCCGC 7800

TACAGGGCGC GTCGCGCCAT TCGCCATTCA GGCTACGCAA CTGTTGGGAA 7850

GGGCGATCGG TGCGGGCCTC TTCGCTATTA CGCCAGCTGG CTGCAGGGGG 7900

GGGGGGGGGG GGGT 7914